RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: _10/500,477A_

Source: _PCT_

Date Processed by STIC: _4/20/06_

ENTERED

PCT

```
 3 <110> APPLICANT: Pettersson, Dan
 4        Fuglsang, Claus Crone
 5        Wu, Wenping
 7 <120> TITLE OF INVENTION: Thermostable Enzyme Compositions
 9 <130> FILE REFERENCE: 10254.204-US
11 <140> CURRENT APPLICATION NUMBER: 10/500,477A
12 <141> CURRENT FILING DATE: 2004-06-29
14 <160> NUMBER OF SEQ ID NOS: 18
16 <170> SOFTWARE: PatentIn version 3.3
18 <210> SEQ ID NO: 1
19 <211> LENGTH: 1008
20 <212> TYPE: DNA
21 <213> ORGANISM: Thermoascus aurantiacus
24 <220> FEATURE:
25 <221> NAME/KEY: sig_peptide
26 <222> LOCATION: (1)..(90)
28 <220> FEATURE:
29 <221> NAME/KEY: CDS
30 <222> LOCATION: (1)..(1005)
32 <220> FEATURE:
33 <221> NAME/KEY: mat_peptide
34 <222> LOCATION: (91)..(1005)
36 <400> SEQUENCE: 1
37 atg aag ctc ggc tct ctc gtg ctc gct ctc agc gca gct agg ctt aca    48
38 Met Lys Leu Gly Ser Leu Val Leu Ala Leu Ser Ala Ala Arg Leu Thr
39 -30             -25                 -20                 -15
41 ctg tcg gcc cct ctc gca gac aga aag cag gag acc aag cgt gcg aaa    96
42 Leu Ser Ala Pro Leu Ala Asp Arg Lys Gln Glu Thr Lys Arg Ala Lys
43             -10                 -5                  -1  1
45 gta ttc caa tgg ttc ggt tcg aac gag tcc ggt gct gaa ttc gga agc   144
46 Val Phe Gln Trp Phe Gly Ser Asn Glu Ser Gly Ala Glu Phe Gly Ser
47         5                   10                  15
49 cag aac ctt cca gga gtc gag gga aag gat tat ata tgg cct gat ccc   192
50 Gln Asn Leu Pro Gly Val Glu Gly Lys Asp Tyr Ile Trp Pro Asp Pro
51     20                  25                  30
53 aac acc att gac aca ttg atc agc aag ggg atg aac atc ttt cgt gtc   240
54 Asn Thr Ile Asp Thr Leu Ile Ser Lys Gly Met Asn Ile Phe Arg Val
55 35                  40                  45                  50
57 ccc ttt atg atg gag aga ttg gtt ccc aac tca atg acc ggc tct ccg   288
58 Pro Phe Met Met Glu Arg Leu Val Pro Asn Ser Met Thr Gly Ser Pro
59             55                  60                  65
61 gat ccg aac tac ctg gca gat ctc ata gcg act gta aat gca atc acc   336
62 Asp Pro Asn Tyr Leu Ala Asp Leu Ile Ala Thr Val Asn Ala Ile Thr
```

```
 63              70              75              80
 65 cag aaa ggt gcc tac gcc gtc gtc gat cct cat aac tac ggc aga tac    384
 66 Gln Lys Gly Ala Tyr Ala Val Val Asp Pro His Asn Tyr Gly Arg Tyr
 67           85              90              95
 69 tac aat tct ata atc tcg agc cct tcc gat ttc cag acc ttc tgg aaa    432
 70 Tyr Asn Ser Ile Ile Ser Ser Pro Ser Asp Phe Gln Thr Phe Trp Lys
 71       100             105             110
 73 acg gtc gcc tca cag ttt gct tcg aat cca ctg gtc atc ttc gac act    480
 74 Thr Val Ala Ser Gln Phe Ala Ser Asn Pro Leu Val Ile Phe Asp Thr
 75 115             120             125             130
 77 aat aac gaa tac cac gat atg gac cag acc tta gtc ctc aat ctc aac    528
 78 Asn Asn Glu Tyr His Asp Met Asp Gln Thr Leu Val Leu Asn Leu Asn
 79           135             140             145
 81 cag gcc gct atc gac ggc atc cgt tcc gcc gga gcc act tcc cag tac    576
 82 Gln Ala Ala Ile Asp Gly Ile Arg Ser Ala Gly Ala Thr Ser Gln Tyr
 83           150             155             160
 85 atc ttt gtc gag ggc aat tcg tgg acc ggg gca tgg acc tgg acg aac    624
 86 Ile Phe Val Glu Gly Asn Ser Trp Thr Gly Ala Trp Thr Trp Thr Asn
 87           165             170             175
 89 gtg aac gat aac atg aaa agc ctg acc gac cca tct gac aag atc ata    672
 90 Val Asn Asp Asn Met Lys Ser Leu Thr Asp Pro Ser Asp Lys Ile Ile
 91       180             185             190
 93 tac gag atg cac cag tac ctg gac tct gac gga tcc ggg aca tca gcg    720
 94 Tyr Glu Met His Gln Tyr Leu Asp Ser Asp Gly Ser Gly Thr Ser Ala
 95 195             200             205             210
 97 acc tgc gta tct tcg acc atc ggt caa gag cga atc acc agc gca acg    768
 98 Thr Cys Val Ser Ser Thr Ile Gly Gln Glu Arg Ile Thr Ser Ala Thr
 99           215             220             225
101 cag tgg ctc agg gcc aac ggg aag aag ggc atc atc ggc gag ttt gcg    816
102 Gln Trp Leu Arg Ala Asn Gly Lys Lys Gly Ile Ile Gly Glu Phe Ala
103           230             235             240
105 ggc gga gcc aac gac gtc tgc gag acg gcc atc acg ggc atg ctg gac    864
106 Gly Gly Ala Asn Asp Val Cys Glu Thr Ala Ile Thr Gly Met Leu Asp
107           245             250             255
109 tac atg gcc cag aac aca gac gtc tgg act ggc gcc atc tgg tgg gcg    912
110 Tyr Met Ala Gln Asn Thr Asp Val Trp Thr Gly Ala Ile Trp Trp Ala
111       260             265             270
113 gcc ggg ccg tgg tgg gga gac tac ata ttc tcc atg gag ccg gac aat    960
114 Ala Gly Pro Trp Trp Gly Asp Tyr Ile Phe Ser Met Glu Pro Asp Asn
115 275             280             285             290
117 ggc atc gcg tat cag cag ata ctt cct att ttg act ccg tat ctt tga   1008
118 Gly Ile Ala Tyr Gln Gln Ile Leu Pro Ile Leu Thr Pro Tyr Leu
119           295             300             305
122 <210> SEQ ID NO: 2
123 <211> LENGTH: 335
124 <212> TYPE: PRT
125 <213> ORGANISM: Thermoascus aurantiacus
127 <400> SEQUENCE: 2
129 Met Lys Leu Gly Ser Leu Val Leu Ala Leu Ser Ala Ala Arg Leu Thr
```

Input Set : A:\PTO.AMC.txt
Output Set: N:\CRF4\04202006\J500477A.raw

```
130 -30                  -25                  -20                  -15
133 Leu Ser Ala Pro Leu Ala Asp Arg Lys Gln Glu Thr Lys Arg Ala Lys
134                 -10                  -5              -1  1
137 Val Phe Gln Trp Phe Gly Ser Asn Glu Ser Gly Ala Glu Phe Gly Ser
138         5                   10                  15
141 Gln Asn Leu Pro Gly Val Glu Gly Lys Asp Tyr Ile Trp Pro Asp Pro
142     20                  25                  30
145 Asn Thr Ile Asp Thr Leu Ile Ser Lys Gly Met Asn Ile Phe Arg Val
146 35                  40                  45                  50
149 Pro Phe Met Met Glu Arg Leu Val Pro Asn Ser Met Thr Gly Ser Pro
150             55                  60                  65
153 Asp Pro Asn Tyr Leu Ala Asp Leu Ile Ala Thr Val Asn Ala Ile Thr
154         70                  75                  80
157 Gln Lys Gly Ala Tyr Ala Val Val Asp Pro His Asn Tyr Gly Arg Tyr
158         85                  90                  95
161 Tyr Asn Ser Ile Ile Ser Ser Pro Ser Asp Phe Gln Thr Phe Trp Lys
162     100                 105                 110
165 Thr Val Ala Ser Gln Phe Ala Ser Asn Pro Leu Val Ile Phe Asp Thr
166 115                 120                 125                 130
169 Asn Asn Glu Tyr His Asp Met Asp Gln Thr Leu Val Leu Asn Leu Asn
170             135                 140                 145
173 Gln Ala Ala Ile Asp Gly Ile Arg Ser Ala Gly Ala Thr Ser Gln Tyr
174         150                 155                 160
177 Ile Phe Val Glu Gly Asn Ser Trp Thr Gly Ala Trp Thr Trp Thr Asn
178         165                 170                 175
181 Val Asn Asp Asn Met Lys Ser Leu Thr Asp Pro Ser Asp Lys Ile Ile
182     180                 185                 190
185 Tyr Glu Met His Gln Tyr Leu Asp Ser Asp Gly Ser Gly Thr Ser Ala
186 195                 200                 205                 210
189 Thr Cys Val Ser Ser Thr Ile Gly Gln Glu Arg Ile Thr Ser Ala Thr
190             215                 220                 225
193 Gln Trp Leu Arg Ala Asn Gly Lys Lys Gly Ile Ile Gly Glu Phe Ala
194         230                 235                 240
197 Gly Gly Ala Asn Asp Val Cys Glu Thr Ala Ile Thr Gly Met Leu Asp
198         245                 250                 255
201 Tyr Met Ala Gln Asn Thr Asp Val Trp Thr Gly Ala Ile Trp Trp Ala
202     260                 265                 270
205 Ala Gly Pro Trp Trp Gly Asp Tyr Ile Phe Ser Met Glu Pro Asp Asn
206 275                 280                 285                 290
209 Gly Ile Ala Tyr Gln Gln Ile Leu Pro Ile Leu Thr Pro Tyr Leu
210             295                 300                 305
```

213 <210> SEQ ID NO: 3
214 <211> LENGTH: 21
215 <212> TYPE: PRT
216 <213> ORGANISM: Thermoascus aurantiacus
219 <220> FEATURE:
220 <221> NAME/KEY: MISC_FEATURE
221 <223> OTHER INFORMATION: N-terminal peptide
223 <220> FEATURE:

RAW SEQUENCE LISTING   DATE: 04/20/2006
PATENT APPLICATION: US/10/500,477A   TIME: 12:11:49

Input Set : A:\PTO.AMC.txt
Output Set: N:\CRF4\04202006\J500477A.raw

```
     224 <221> NAME/KEY: MISC_FEATURE
     225 <222> LOCATION: (2)..(2)
     226 <223> OTHER INFORMATION: Xaa in position 2 means any amino acid
     228 <400> SEQUENCE: 3
W--> 230 Asn Xaa Leu Val Phe Thr Ser Phe Gly Ser Asn Glu Ser Gly Ala Glu
     231 1               5                   10                  15
     234 Phe Gly Ser Gln Asn
     235               20
     238 <210> SEQ ID NO: 4
     239 <211> LENGTH: 20
     240 <212> TYPE: DNA
     241 <213> ORGANISM: Artificial
     243 <220> FEATURE:
     244 <223> OTHER INFORMATION: Primer
     247 <220> FEATURE:
     248 <221> NAME/KEY: misc_feature
     249 <222> LOCATION: (9)..(9)
     250 <223> OTHER INFORMATION: n is a, c, g, or t
     252 <220> FEATURE:
     253 <221> NAME/KEY: misc_feature
     254 <222> LOCATION: (12)..(12)
     255 <223> OTHER INFORMATION: n is a, c, g, or t
     257 <220> FEATURE:
     258 <221> NAME/KEY: misc_feature
     259 <222> LOCATION: (15)..(15)
     260 <223> OTHER INFORMATION: n is a, c, g, or t
     262 <400> SEQUENCE: 4
W--> 263 aaygartcng gngcngaatt                                    20
     266 <210> SEQ ID NO: 5
     267 <211> LENGTH: 20
     268 <212> TYPE: DNA
     269 <213> ORGANISM: Artificial
     271 <220> FEATURE:
     272 <223> OTHER INFORMATION: Primer
     275 <220> FEATURE:
     276 <221> NAME/KEY: misc_feature
     277 <222> LOCATION: (9)..(9)
     278 <223> OTHER INFORMATION: n is a, c, g, or t
     280 <220> FEATURE:
     281 <221> NAME/KEY: misc_feature
     282 <222> LOCATION: (12)..(12)
     283 <223> OTHER INFORMATION: n is a, c, g, or t
     285 <220> FEATURE:
     286 <221> NAME/KEY: misc_feature
     287 <222> LOCATION: (15)..(15)
     288 <223> OTHER INFORMATION: n is a, c, g, or t
     290 <400> SEQUENCE: 5
W--> 291 aaygartcng gngcngagtt                                    20
     294 <210> SEQ ID NO: 6
```

●   ●

```
      295 <211> LENGTH: 20
      296 <212> TYPE: DNA
      297 <213> ORGANISM: Artificial
      299 <220> FEATURE:
      300 <223> OTHER INFORMATION: Primer
      303 <220> FEATURE:
      304 <221> NAME/KEY: misc_feature
      305 <222> LOCATION: (12)..(12)
      306 <223> OTHER INFORMATION: n is a, c, g, or t
      308 <220> FEATURE:
      309 <221> NAME/KEY: misc_feature
      310 <222> LOCATION: (15)..(15)
      311 <223> OTHER INFORMATION: n is a, c, g, or t
      313 <400> SEQUENCE: 6
W--> 314 aaygaragkg gngcngaatt                                       20
      317 <210> SEQ ID NO: 7
      318 <211> LENGTH: 20
      319 <212> TYPE: DNA
      320 <213> ORGANISM: Artificial
      322 <220> FEATURE:
      323 <223> OTHER INFORMATION: Primer
      326 <220> FEATURE:
      327 <221> NAME/KEY: misc_feature
      328 <222> LOCATION: (12)..(12)
      329 <223> OTHER INFORMATION: n is a, c, g, or t
      331 <220> FEATURE:
      332 <221> NAME/KEY: misc_feature
      333 <222> LOCATION: (15)..(15)
      334 <223> OTHER INFORMATION: n is a, c, g, or t
      336 <400> SEQUENCE: 7
W--> 337 aaygaragkg gngcngagtt                                       20
      340 <210> SEQ ID NO: 8
      341 <211> LENGTH: 18
      342 <212> TYPE: DNA
      343 <213> ORGANISM: Artificial
      345 <220> FEATURE:
      346 <223> OTHER INFORMATION: Primer
      348 <400> SEQUENCE: 8
      349 aagatgtact gggaagtg                                         18
      352 <210> SEQ ID NO: 9
      353 <211> LENGTH: 21
      354 <212> TYPE: DNA
      355 <213> ORGANISM: Artificial
      357 <220> FEATURE:
      358 <223> OTHER INFORMATION: Primer
      360 <400> SEQUENCE: 9
      361 tggttgagat tgaggactaa g                                     21
      364 <210> SEQ ID NO: 10
      365 <211> LENGTH: 21
```

RAW SEQUENCE LISTING ERROR SUMMARY   DATE: 04/20/2006
PATENT APPLICATION: US/10/500,477A   TIME: 12:11:50

Input Set : **A:\PTO.AMC.txt**
Output Set: **N:\CRF4\04202006\J500477A.raw**

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:3; Xaa Pos. 2
Seq#:4; N Pos. 9,12,15
Seq#:5; N Pos. 9,12,15
Seq#:6; N Pos. 12,15
Seq#:7; N Pos. 12,15

Invalid <213> Response:
Use of "Artificial" only as "<213> Organism" response is incomplete,
per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:4,5,6,7,8,9,10,11,12,13

VERIFICATION SUMMARY   DATE: 04/20/2006
PATENT APPLICATION: **US/10/500,477A**   TIME: 12:11:50

Input Set : **A:\PTO.AMC.txt**
Output Set: **N:\CRF4\04202006\J500477A.raw**

```
L:230 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3 after pos.:0
L:263 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4 after pos.:0
L:291 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:5 after pos.:0
L:314 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:6 after pos.:0
L:337 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:0
```